The Behavioural Effect of Electronic Home Energy Reports: Evidence from a Randomised Field Trial in the United States

Marisa L. Henry[a], Paul J. Ferraro[b], Andreas Kontoleon[c,*]

[a]*Department of Environmental Health and Engineering, Whiting School of Engineering, Johns Hopkins University, 3400 N Charles St, Baltimore, MD 21218*
[b]*Carey Business School and the Department of Environmental Health and Engineering, a joint department of the Bloomberg School of Public Health and Whiting School of Engineering, Johns Hopkins University, 100 International Drive, Baltimore, MD 21202*
[c]*Department of Land Economy, University of Cambridge, 19 Silver Street, Cambridge CB3 9EP, UK*

Abstract

Behavioural interventions, such as informational nudges, have become an increasingly popular strategy in demand-side energy management. In particular, home energy reports (HERs) have been used to induce behavioural change among residential consumers. These HERs typically provide peer comparisons of energy use and information about energy savings opportunities. Despite the growing prevalence of HERs and a shift from postal to electronic delivery of HERs, the experimental evidence base of their effectiveness comes primarily from HERs delivered by post from a single vendor (Opower). Whether that evidence generalises to other programmes and to the electronic delivery of HERs is unclear. This paper reports new evidence for HER effectiveness from a 12-month field experiment with approximately 9,000 households that tested electronic HER programme in a deregulated American residential electricity market. Despite high non-compliance with HER delivery, the programme reduced household electricity consumption by 2.9%, 95% CI [-5.0%, -0.76%]. This estimated reduction is consistent with prior estimated impacts of HERs delivered by post and implies electronic HERs are at least as effective as reports delivered by post in reducing elec-

*Corresponding author
Email address: `ak219@cam.ac.uk` (Andreas Kontoleon)

tricity consumption, while they are administered at a lower cost.

Keywords: Behaviour Change, Impact Evaluation, Energy Conservation, Nudge

---

Highlights

- Electronic home energy reports reduce electricity consumption by 2.9%
- Electronic home energy reports appear at least as effective as physical reports
- Electronic home energy report delivery is more cost effective than postal delivery

1. Introduction

Widespread concerns about energy reliability and anthropogenic climate change have led to growing pressures on the energy sector to manage energy demands (Dietz et al., 2009). To encourage demand-side energy reductions, both private and public sector actors have increasingly relied on behavioural programmes (Karlin et al., 2015). Behavioural programmes apply social science theories of human behaviour and decision making to encourage behaviour change without eliminating choice or changing economic incentives (Sussman and Chikumbo, 2016). The impacts of such programmes, often called "nudges", have been quantified through field experiments in a variety of economic sectors (Thaler and Sunstein, 2008).

In the energy sector, the most extensive experimental evidence for the effectiveness of behavioural programmes comes from the home energy report (HER) programmes run by the company Opower (now part of Oracle). In collaboration with dozens of energy utilities across the United States, Opower ran a series of field experiments to evaluate the effect of HERs on residential energy consumption. Based on research suggesting social comparisons influence energy consumption, HERs use social comparisons and personalised information provision to promote behaviour change (Nolan et al., 2008; Schultz et al., 2007). The reports give individuals information on how their electricity use compares with households of similar size in their area. This information is combined with text implying lower energy consumption is socially "good" and with personalised strategies for lowering energy consumption, which are based on the household's energy use profile. Allcott (2011) first estimated the average treatment effect of HERs using data from approximately 600,000 urban and rural households that were randomised as part of the Opower programmes.

The Opower programmes reduced energy consumption by an estimated 1.4% to 3.3%, depending on the regional market (Allcott, 2011). These effects are reported to persist in the long run due to improved capital stock of more efficient technologies and changes in behavioural habits (Allcott and Rogers, 2014; Brandon et al., 2017; Ayres et al., 2009). Based on these results, HERs are cost-effective compared to energy efficiency programmes (Allcott and Mullainathan, 2010). Similarly, personalised, pro-social messaging has

Abbreviations: HER, home energy report; REP, retail electricity provider; ITT, intent-to-treat been shown to be cost-effective in encouraging reductions in residential water consumption, indicating analogous behavioural programmes can be applied successfully in other sectors (Ferraro and Price, 2013; Bernedo et al., 2014).

Since the evaluation of Opower's programmes, utilities and other third-party companies have developed and implemented their own variations of home energy reports. However, to our knowledge, only one of these variations has been formally evaluated with experimental designs in the peer-reviewed literature. Byrne et al. (2018) studies an Australian programme that uses high-frequency smart meter data to provide social norms and personalised information to households via biweekly emails and a smart meter web portal. However, the HERs are accompanied by feedback from the granular, smart meter data and an interactive web portal, making it difficult to disentangle the effect of the HERs from the other components of the intervention.

This gap in the evidence base is important given that the evaluated Opower programmes have been shown to have partner selection bias; the utilities that use Opower differ systematically from other utilities (Allcott and Mullainathan, 2012). This bias potentially restricts the generalisability of the Opower results. In other words, the estimated average treatment effects reported in the literature may not reflect expected impacts in other HER programmes. Furthermore, the Opower experiments involved physical reports, delivered by post. Yet many HER programmes, including Opower's, have shifted to electronic communication, which tends to have lower read rates than direct mail (DMN, 2012).

Using new data from a randomised control trial conducted over a period of 12 months with nearly 9,000 households, this paper adds to the literature on HERs by estimating the causal effect of an electronic HER programme on residential electricity consumption. Despite high non-compliance with HER delivery, the HER programme is estimated to have reduced monthly electricity consumption by 2.9%, 95% CI [-5.00%, -0.76%], which is consistent with published estimates from the Opower programmes with physical reports. This result suggests electronically delivered HERs are at least as effective in reducing electricity consumption as HERs delivered by post and achieve their impacts at lower cost. The paper also speaks to the broader literature on using social comparisons and informational nudges to influence behaviour in an environmental policy context (Croson and Treich, 2014).

2. Methods and Data

2.1. Experimental Design

In late 2015, a private home sensing and software company began working with a retail electricity provider (REP) to quantify the impact of electronically provided HERs on residential electricity consumption in one American state[1]. At the time of the study, the REP had customers from seven utilities in the state.

The software company implemented a randomised control trial using a sample that comprised all households registered with the REP in the state in November 2015 (N = 9,383). Households were blocked by utility provider, and then 14.5% of the customers within each block were randomly assigned to a control group that would receive status quo communication from the REP. The remaining households from each utility were assigned to the treatment group and received the additional customised HERs developed by the software company. One utility had only 3 customers, all of which were assigned to treatment. This utility is excluded from analysis due to the lack of control group (Figure 1).

The reports included information on the customers' monthly electricity use in kilowatt hours (kWh), a social comparison module comparing the users' electricity consumption with that of similar households in their area, and customised advice on reducing electricity use. The utilities provided the use data in the reports, and the software company's proprietary algorithms generated the customised advice (Cetin et al., 2016). The completed reports were sent to the REP, and the REP distributed them to households via email. The email contained both an attached PDF of the HER and a link to a private web portal, where customers could view their current and past reports. An example HER is shown in Appendix A.

[1]We obtained the data from the home sensing and software company. In order to receive permission to make the de-identified data publicly accessible, we agreed to not name the company or the state in which the experiment was run. However, we note the field experiment was run in a U.S. state in which the electricity market has been deregulated and utility revenues decoupled from electricity volume throughput, removing utility disincentive to pursuing consumer energy efficiency (Nissen and Williams, 2016). Additionally, the states' energy efficiency resources standards add a positive financial incentive for utilities to pursue demand-side energy efficiency programmes (ACEEE, 2017). These incentives may encourage retail electricity providers in the state to pursue behavioural energy efficiency programmes such as home energy reports.

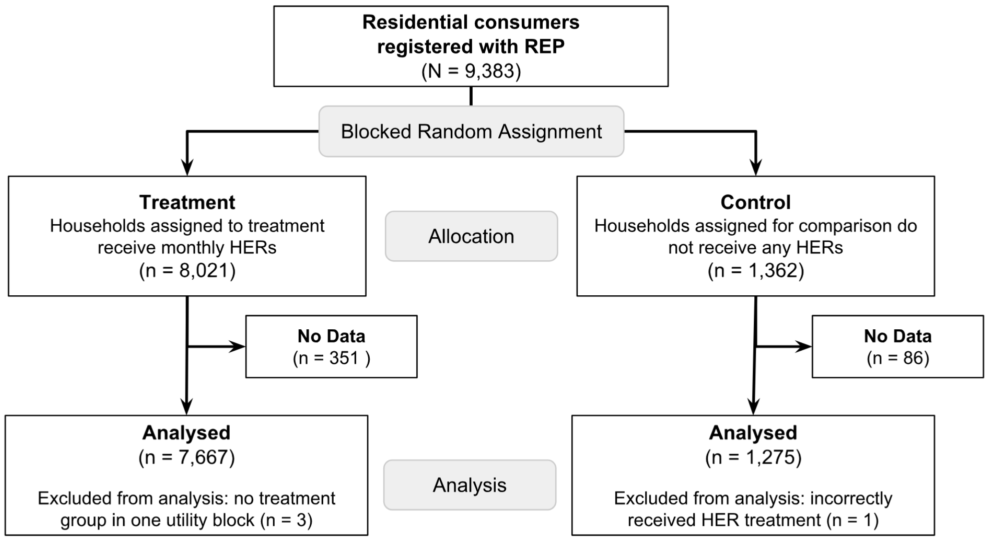

Figure 1: Flow diagram of the progress through the randomised control trial.

2.2. Data and Baseline Characteristics

Household electricity use and HER delivery are available at monthly intervals over the course of the study period. Monthly baseline electricity use data is available for each month since the household joined the REP, which varies between households. In December 2015, prior to the HER intervention, the average logged monthly electricity use in treatment and control groups is statistically equivalent, both within utility and overall (Table 1).

The de-identified data and code are available through the Open Science Framework at https://osf.io/a2fhq/.

2.3. Attrition

Randomisation of the HER treatment took place in November 2015. The first HERs were sent in January 2016, reporting electricity use from the previous month. Our post-intervention study period is twelve months (January 2016 through December 2016), which is substantially longer than most experimental studies on information-based energy conservation experiments. In a meta-analysis of 156 published experiments from 1975 to 2012, Delmas et al.

Table 1: Average monthly electricity use in December 2015, the month before the first HERs were sent to treatment households. Electricity use is compared between treatment assignment using a t-test of average log electricity use in December 2015 and p-values are shown in parentheses.

	Treatment		**Control**		**Difference**
Utility	**n**	**Use (kWh)**	**n**	**Use (kWh)**	**(p-value)**
1	3,931	969.00	654	967.26	1.74 (0.48)
2	2,131	876.29	355	1001.95	-125.66 (0.35)
3	1,251	738.29	209	747.17	-8.88 (0.33)
4	317	882.15	50	894.04	-11.89 (0.74)
5	31	791.76	6	534.75	257.01 (0.40)
6	6	367.00	1	NA	NA
All	7,667	901.40	1,275	933.11	-31.71 (0.75)

(2013) reported that about 60 percent of the studies lasted for three months or less, and they called for studies of longer duration.

One issue that arises in longer studies is attrition. In the month after randomisation, 4.7% of households left the REP, and the utilities transferred no use data on these customers. As expected, given the HERs had not yet been sent, the proportion of households without data is similar between treatment (4.4%) and control (6.4%) groups, and thus their exclusion from the analysis causes no potential problems for estimation.

Attrition of customers continued throughout the treatment period at about 450 households per month, and the monthly rate is similar between the treatment and control groups (Figure 2). Over the entire twelve month period after the intervention began, households assigned to treatment have an attrition rate of 55.5% compared to 54.0% in the control group, a difference of 1.5%, 95%CI [-1.39%, 4.52%]. Furthermore, the average pre-intervention electricity use of households who attrite is similar between treatment and control groups: the difference between them is 0.025 log(kWh/month), 95%CI [-0.0002, 0.048].

If the rate of attrition is independent of potential electricity use with and without HERs, attrition does not create bias in the estimation of the treatment effect. Although no direct test can confirm or reject this independence assumption, the evidence from Figure 2, the comparison of pre-treatment use, and the overall difference in rates is consistent with the assumption.

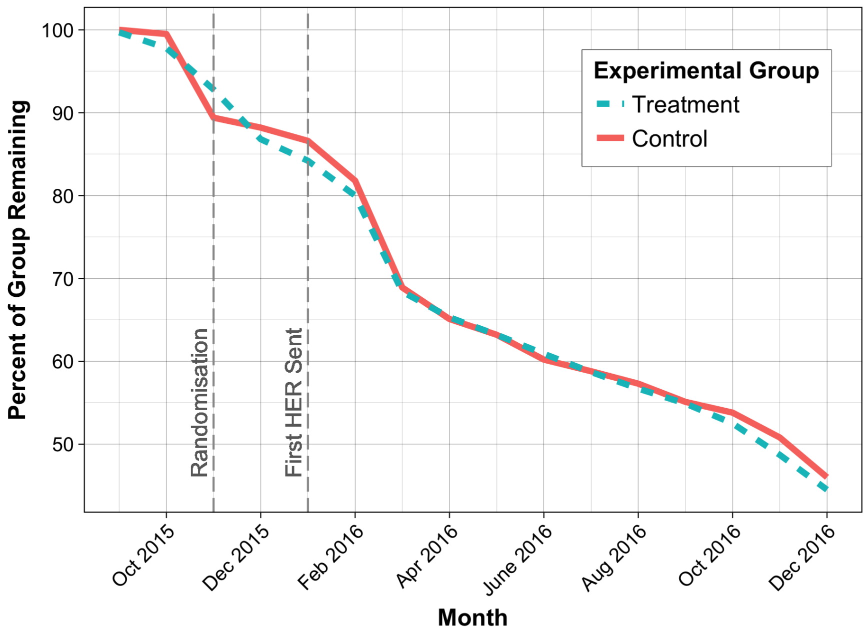

Figure 2: Percent of each experimental group for which electricity use is observed over the study period. Exact values are provided in Appendix B.

.

2.4. Intent-to-treat Effect

We focus on estimating the intent-to-treat (ITT) effect of the electronically delivered HER reports. The ITT is the effect of being assigned to receive an HER, regardless of whether the HER was delivered or read by the household. Randomisation of the HER delivery ensures the ITT can be estimated without bias, even when there is experimental non-compliance, i.e. when households did not receive some monthly reports[2]. Furthermore, the ITT estimate is the estimand of interest for the REP and policy makers

[2]Simply excluding the non-complying households from the treatment group would create bias in our estimator if the factors that affect whether a household is sent a report are also correlated with household electricity use

who are interested in the expected impact of a real-world, electronic HER programme, not an idealised programme that has zero non-compliance.

The programme aimed to send reports to households on a monthly basis for a period of 12 months. However, because of periodic delays in utilities reading electricity meters and transferring the use data between the utilities, REP, and the company creating the HERs, most households did not receive an HER every month. A relatively large percentage (36 percent) of households assigned to treatment did not receive any HERs; in their case, the timing of meter readings and data transfer never matched the timing requirements for creating a monthly HER for them. Although the ITT can be estimated without bias from non-compliance, we note the average logged electricity use of those who were assigned to treatment but did not receive any reports is statistically equivalent to the average logged electricity use for those who received at least one report. This suggests that those who did not receive reports did not have systematically higher or lower electricity consumption relative to the rest of the treatment group. The frequency of reports sent to treatment group households is reported in Table 2.

Table 2: Number of reports received by households assigned to the treatment group.

Number of Reports Received	**Number of Households**	**Percent of treatment group**
0	2,757	35.98%
1-3	2,314	30.17%
4-6	508	6.62%
7-9	618	8.06%
10-12	1,470	19.17%

The primary model specification relates log electricity use with treatment assignment and is of the form:

$$ln(Y_{it}) = \beta_0 + \beta_1 \cdot trt_{it} + U_i + \mu_{ym} + \alpha_i + \epsilon_{it} \tag{1}$$

where Y_{it} is the monthly electricity use of household i in month t in units of kWh per month; trt_{it} is the binary treatment assignment of each household during each month; U_i is a series of indicators for the utility of household i; μ_{ym} is a series of month-by-year dummy indicators; α_i represents a household error term; and ϵ_{it} is random error. Logs were used to adjust for the right-skew in the electricity use values and so model outputs can be interpreted as a percentage change between control and treatment. This empirical strategy is a standard model for identifying intent to treat effects when households are randomly assigned to treatment (Gerber and Green, 2012).

The primary parameter of interest in this specification is β_1, which indicates the ITT: the effect of assignment to the treatment group on electricity consumption. The model parameters are estimated with generalised least squares using a random effects panel data model. Standard errors are clustered by household to account for serial autocorrelation.

3. Results and Discussion

3.1. Estimated ITT

Running a random effects panel regression of the form in Equation 1, the intent-to-treat effect, β_1, is estimated as -2.88% (SE = 1.083%) with a 95% confidence interval of -5.0% to -0.76%. If we remove households with monthly electricity use values in the top one percent of all observed use (>3,201 kWh/month), the ITT is -2.60%, 95% CI [-4.79, -0.42], indicating the effect is robust to removing these "outliers".

Given that some households were not compliant with their treatment assignment (they received no HER), we cannot estimate the average treatment effect of being sent an HER unless we were willing to assume that non-compliance was independent of potential energy use, an assumption that is unlikely to be true. Nevertheless, we can estimate the causal effect of being sent at least one HER for the subgroup of households that comply with their treatment assignment. This complier subgroup is comprised of households who get at least one HER when they are assigned to treatment and do not get an HER when they are not assigned to treatment. To estimate this causal effect, we use a two-stage panel data regression estimator with randomized assignment as an instrumental variable for receiving at least one HER. We estimate that this complier average causal effect of receiving at least one HER is -4.84%, 95% CI [-8.40, -1.28], an effect that is more than 1.5 times the estimated ITT effect.

The estimated intent-to-treat effect of -2.88% is consistent with the upper end of the average treatment effect estimates from Allcott's (2011) assessment of Opower's HER programmes on energy use across the United States. Assuming that HERs do not cause an increase in electricity consumption in any household, the intent-to-treat effect represents a lower bound on the average treatment effect of this programme. This suggests that if compliance was increased, the causal effect of the programme on electricity consumption would be larger.

3.2. Heterogeneous Treatment Effects

Understanding how the treatment effect varies conditional on observable characteristics of the households can help better target the intervention to subgroups who would be most responsive (i.e., improve programme cost-effectiveness) and can shed light on the underlying mechanisms through which the intervention affects behaviour (Ferraro and Miranda, 2013). The moderating variable that is most frequently examined in prior studies is pre-treatment energy use. REPs are usually able to observe this variable, and thus could condition on it when targeting an HER intervention. Allcott (2011) reported that higher energy users before the HER programme conserved more energy after receiving HERs. Similar results have been reported in the context of water use (Ferraro and Price, 2013; Ferraro and Miranda, 2013).

To test for heterogeneous treatment effects conditional on pre-treatment electricity use in our study, we define high users as individuals whose energy use was above the median in the month prior to when the first HERs were sent. We use only one month of pre-treatment use because the extent of pre-treatment data available varies between households. We add this indicator variable to the regression model in Equation (1) both on its own and interacted with the treatment variable as:

$$ln(Y_{it}) = \beta_0 + \beta_1 \cdot trt_{it} + \beta_2 \cdot trt_{it} \cdot H_i + H_i + U_i + \mu_{ym} + \alpha_i + \epsilon_{it} \qquad (2)$$

where H_i is the indicator for users with above median energy consumption the month before treatment.

In this specification, β_1 (the estimated coefficient on the treatment variable) is the estimated ITT for the low users, and $\beta_1 + \beta_2$ (the combined coefficients of the treatment variable and the treatment variable interacted with the indicator variable for high user) is the ITT for the high users. This subgroup definition was used by Ferraro and Price (2013) and is coarser but similar to the definition by decile used by Allcott (2011). The estimated coefficient on the treatment variable is -3.01% (CI 95%[-5.41%, -0.62%]), which is essentially unchanged from the original ITT estimate. The interaction term's coefficient (β_2) is small and not statistically different from zero: 0.59% (CI 95%[-1.59%, 2.78%]). Thus, in contrast to previous studies, we do not detect any difference in the behavioural responses of low and high energy users.

4. Conclusions and Policy Implications

Programme administrators have a number of levers at their disposal to influence residential energy use, including traditional financial incentives. Increasingly, however, they are turning to behavioural mechanisms. Based on theories from the behavioural sciences, behavioural interventions are often relatively simple to implement and have been shown to be cost-effective (Allcott and Rogers, 2014). In the energy sector, home energy reports use social comparisons and personalised information to encourage residential energy use reductions. Prior published evidence on the impacts of these reports come from reports delivered by post from a single vendor (Opower). However, today there are many other utilities and third parties who have developed their own versions of home energy reports, and many of these reports are delivered electronically.

Using data from a randomised field experiment aimed at reducing residential electricity use, this analysis estimates the effect of home energy reports developed by a private home sensing and software company and delivered electronically. The results indicate the programme caused a 2.9% reduction in residential electricity consumption. Despite non-compliance in report delivery, the estimated impact lies on the upper end of previous estimated impacts. Whether these results are generalisable to contexts outside of this study's geographic context is an open empirical question. The customer base, management styles, and regulatory contexts of electricity providers that elect to participate in home energy report programmes may differ from utilities that do not participate, reducing the external validity of the estimated impacts (Allcott and Mullainathan, 2012). Nevertheless, given that the estimated impacts are consistent with prior estimated impacts of HERs delivered by post, they imply that, at least in some contexts, electronically delivered HERs are as effective as physical reports in reducing electricity consumption and are more cost effective.

Appendix A. Example Home Energy Report

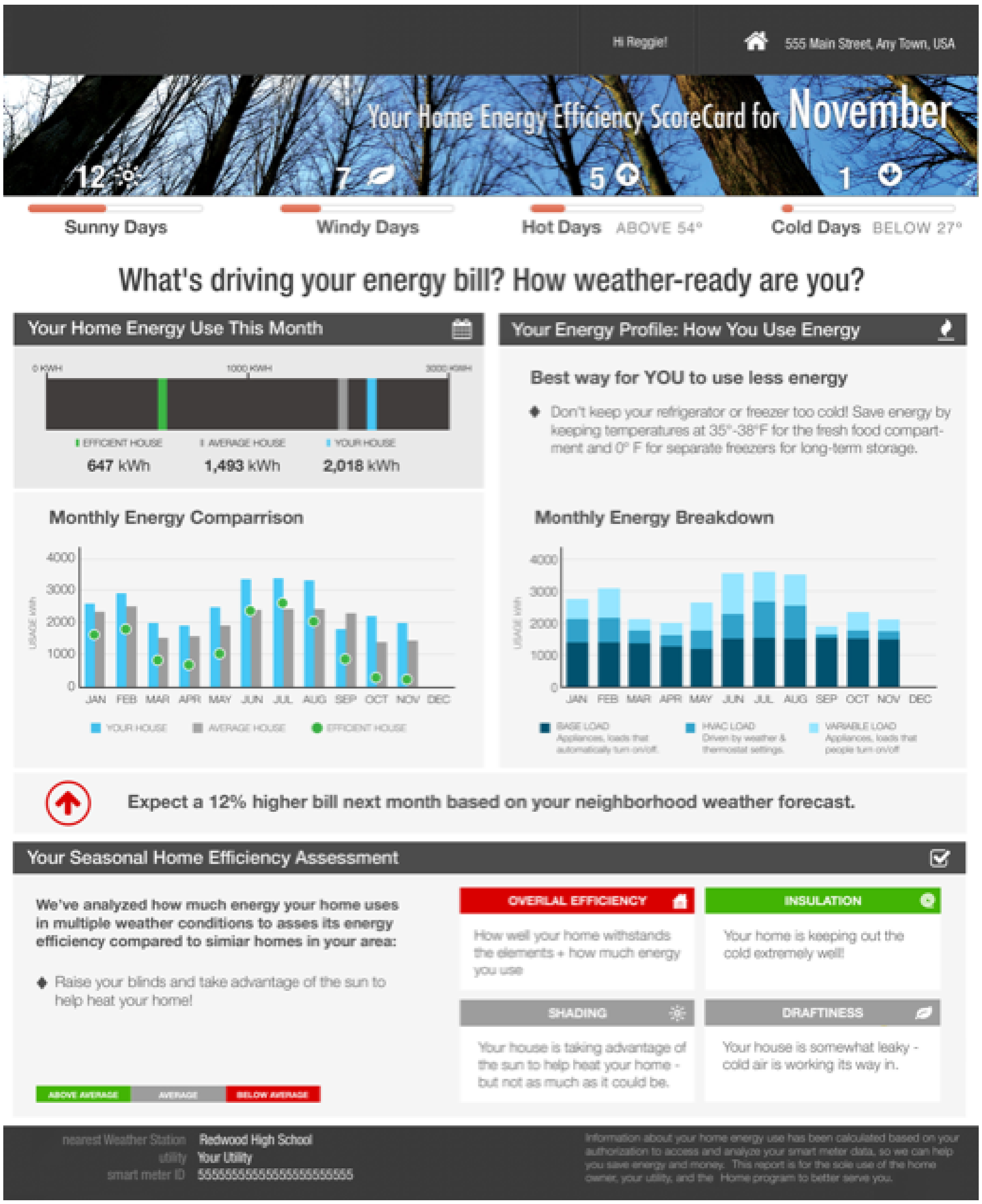

Figure A.3: An example of the layout of the home energy reports used in this study.

Appendix B. Attrition

Table B.3 below shows the cumulative attrition by experimental group for each month of the study period, as plotted in Figure 2.

Table B.3: Cumulative attrition, by experimental group.

Month	Treatment (n=7,667)		Control (n=1,275)	
	Cumulative Attrition	Percent Remaining	Cumulative Attrition	Percent Remaining
October 2015	167	97.8%	6	99.5%
November 2015	554	92.8%	135	89.4%
December 2015	1009	86.8%	150	88.2%
January 2016	1209	84.2%	171	86.6%
February 2016	1531	80.0%	233	81.7%
March 2016	2424	68.4%	397	68.9%
April 2016	2659	65.3%	446	65.0%
May 2016	2825	63.1%	470	63.1%
June 2016	2994	60.9%	509	60.1%
July 2016	3167	58.7%	526	58.7%
August 2016	3316	56.7%	545	57.3%
September 2016	3455	54.9%	573	55.1%
October 2016	3643	52.5%	590	53.7%
November 2016	3930	48.7%	628	50.7%
December 2016	4257	44.5%	689	46.0%

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data statement

The dataset used in this analysis was shared by the a private software company. In order to share the data, we agreed to not name the company or the state in which the experiment was run. The de-identified data and code are available through the Open Science Framework at https://osf.io/a2fhq/.

References

ACEEE, 2017. State Energy Efficiency Resource Standards (EERS). URL https://aceee.org/sites/default/files/state-eers-0117.pdf

Allcott, H., 2011. Social Norms and Energy Conservation. Journal of Public Economics 95, 1082–1095.

Allcott, H., Mullainathan, S., 2010. Behavior and Energy Policy. Science 327, 1204–1205.

Allcott, H., Mullainathan, S., 2012. External Validity and Partner Selection Bias. NBER Working Paper, 1–44.

Allcott, H., Rogers, T., 2014. The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation. American Economic Review 104, 3003–3037.

Ayres, I., Raseman, S., Shih, A., 2009. Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage. NBER Working Paper No. 15386.

Bernedo, M., Ferraro, P. J., Price, M., 2014. The Persistent Impacts of Norm-Based Messaging and Their Implications for Water Conservation. Journal of Consumer Policy 37, 437–452.

Brandon, A., Ferraro, P., List, J., Metcalfe, R., Price, M., Rundhammer, F., 2017. Do The Effects of Social Nudges Persist? Theory and Evidence from 38 Natural Field Experiments. NBER Working Paper.

Byrne, D. P., Nauze, A. L., Martin, L. A., 2018. Tell Me Something I Don't Already Know: Informedness and the Impact of Information Programs. Review of Economics and Statistics 100, 510–527.

Cetin, K. S., Siemann, M., Sloop, C., 2016. Disaggregation and Future Prediction of Monthly Residential Building Energy Use Data Using Localized Weather Data Network. ACEEE Summer Study on Energy Efficient Buildings.

Croson, R., Treich, N., 2014. Behavioral Environmental Economics: Promises and Challenges. Environmental and Resource Economics 58 (3), 335–351.

Delmas, M. A., Fischlein, M., Asensio, O. I., 2013. Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012. Energy Policy 61, 729–739.
URL http://dx.doi.org/10.1016/j.enpol.2013.05.109

Dietz, T., Gardner, G. T., Gilligan, J., Stern, P. C., Vandenbergh, M. P., 2009. Household Actions can Provide a Behavioral Wedge to Rapidly Reduce US Carbon Emissions. Proceedings of the National Academy of Sciences 106, 18452–18456.

DMN, 2012. Direct Mail Response Rates Beat Digital.
URL https://www.dmnews.com/channel-marketing/direct-mail/news/13059655/dma-direct-mail-response-rates-beat-digital

Ferraro, P. J., Miranda, J. J., 2013. Heterogeneous treatment effects and mechanisms in information-based environmental policies : Evidence from a large-scale field experiment. Resource and Energy Economics 35, 356–379.

Ferraro, P. J., Price, M. K., 2013. Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiment. Review of Economics and Statistics 95, 64–73.

Gerber, A. S., Green, D. P., 2012. Field Experiments: Design, Analysis, and Interpretation. W. W. Norton.

Karlin, B., Zinger, J. F., Ford, R., 2015. The Effects of Feedback on Energy Conservation: A Meta-Analysis. Psychological Bulletin 141, 1205–1227.

Nissen, W., Williams, S., 2016. The Link Between Decoupling and Success in Utility-Led Energy Efficiency. The Electricity Journal 29, 59–65.

Nolan, J. M., Schultz, P. W., Cialdini, R. B., Goldstein, N. J., Griskevicius, V., 2008. Normative Social Influence is Underdetected. Personality and Social Psychology Bulletin 34 (7), 913–923.

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., Griskevicius, V., 2007. Destructive and Reconstructive Power of Social Norms. Psychological Science 18, 429–434.

Sussman, R., Chikumbo, M., 2016. Behavior Change Programs: Status and Impact. ACEEE Report.

Thaler, R., Sunstein, C., 2008. Nudge: Improving Decisions about Health, Wealth, and Happiness. Yale University Press.